Economy through a lens

Bowden, James; Kwiatkowski, Andrzej; Rambaccussing, Dooruj

Published in:
Journal of Comparative Economics

DOI:
10.1016/j.jce.2019.07.002

Publication date:
2019

Economy through a lens: Distortions of Policy Coverage in UK National Newspapers[1]

Abstract: *This paper shows that communication of economic news varies across newspapers in the United Kingdom. We develop new time series of economic news tonality using a unique dataset of policy influenced articles published in major UK newspapers. We show that the volume and tonality of news respond to current economic conditions. For example, the nature of news changes around events of economic uncertainty such as the global financial crisis and the post-EU referendum periods. We also provide illustrative evidence that communication differs across newspaper formats. Tabloids, as opposed to quality newspapers, tend to express news more negatively, and mostly report policy-related news during periods of economic stress. The integral importance of these results is illustrated by news reaction curves showing a strong positive relationship mostly lasting three months between consumer sentiments and news.*

Keywords: Textual Econometrics, Media, Economic Sentiment, Tonality, Economic News, Economic Policy, United Kingdom, Newspapers

[1] The authors are thankful to the four referees and the journal Editor for their valuable and constructive comments.

1. Introduction

The information set from which agents form expectations regarding future economic conditions is influenced by media outlets, especially TV broadcasts and newspapers (Blinder and Krueger, 2004). These forms of communication report experts' forecasts and political perspectives at particularly low cost. In the United Kingdom (UK), newspapers are a significant supplier of economic news. However, the manner in which economic news is reported and consumed can differ significantly in terms of the frequency and format of news (for example, tabloids versus quality newspapers), the political affiliation of the publication, and the socioeconomic and demographic status of its readership. Economic news may therefore be communicated heterogeneously across newspapers, impacting agents in different ways depending on their news source.

Communication of economic news can take various forms. Firstly, newspapers can report on the stance of business cycles; communicating latest numbers on macroeconomic indicators such as economic growth, inflation, wages and unemployment. Secondly, they may report on future economic prospects provided by experts and surveys. Finally, they may also report on economic policies by the central bank or the government. One of the issues when interpreting data in each case is the potential for newspapers to report information in a biased way. In Hamilton (2004), news is a commodity to be consumed rather than a mirror of reality, which makes it in the interest of media companies to produce sellable news. The driving theory is that some news may be portrayed in a manner which appeals to a certain readership, and thus maximize profits on behalf of media companies (Gentzkow and Shapiro, 2010; Mullainathan and Shleifer, 2005). Moreover, the manner in which economic news is reported within tabloid and broadsheet publications must be considered. For example, Boykoff (2008) identifies differences in the manner with which climate change is communicated, dependent on the tabloid or broadsheet format of the news source.

The role of mass media in the formation of important events has benefited from long-dated attention in political, economic and finance literature. History bears witness to the impact of media coverage on future outcomes, such as political elections (DellaVigna and Kaplan, 2007; Hetherington, 1996). At the macroeconomic level, the impact of media has been addressed using Bayesian learning models (Lamla and Lein, 2014; Lamla and Maag, 2012). Lamla and Lein (2014) find that the intensity of news reporting helps consumers update their expectations on inflation. More news about inflation leads to higher inflation expectations, which are subsequently self-fulfilling. Lamla and Maag (2012) find that dispersion in the tone of economic news leads to higher volatility in inflation expectations, while Doms and Morin (2004) find that economic news correlates with patterns of fluctuating consumer spending.[2]

The influence of newspapers and their reporting has already been measured in US data, where a measure to capture media influence (intensity, tone, or a combination of both) is used. Unlike qualitative techniques, measures such as 'intensity' and 'tone' allow for quantitative representation of words and articles which is useful in the construction of time series data. 'Intensity' relates to the frequency of certain keywords appearing in newspapers over a given time frame, while 'tone' attempts to measure the emotional content (sentiment) expressed within text. As an example, a measure of tone can simply be the number of positive words versus number of negative words in a text. Both intensity and tone measures are then used in different regression settings as predictor variables. [3]

This paper seeks to investigate heterogeneity in the communication of economic news sentiments amongst UK newspapers by firstly developing a time series of economic sentiments, inspired by policy-relevant economic news. The procedure

[2] The authors count the frequency with which the word 'recession' appears within news articles and find that 'bad' news is negatively related with consumer spending.
[3] Examples in applied practice include the 'R word index' reported in The Economist (2011), which uses the frequency with which the word 'recession' appears in newspapers as a predictor of US recession.

is closely linked to the literature on the development of sentiment indices from media sources. For example, Baker et al. (2016) construct an index of economic uncertainty based on the number of times the word 'uncertainty' appears in well-read newspapers. Tobback et al. (2016) refine this model to address economic uncertainty through utilisation of tone instead of intensity, by considering the application of popular sentiment analysis algorithms such as Support Vector Machine (SVM) and Naïve Bayes. Related studies have been conducted in other academic fields, such as finance. For example, Tetlock (2007) finds evidence of message board postings influencing stock market movements while Brown et al. (2018) show the impact of Twitter activity on changes within betting markets.

In this paper, the heterogeneity hypothesis is tested for thirteen UK newspapers using measures of sentiment (tonality), volume (intensity), and consensus across articles. The articles considered in the study are about economic policy in the United Kingdom, and two sentiment analysis techniques – Naïve Bayes and Dictionary – are used to compute sentiment scores for each news article on a positive to negative scale. We find that news media sentiment becomes notably less positive during the global financial crisis; our newspaper sentiment measures decreased by 63.4% (Naïve Bayes) and 36.2% (Dictionary) from the baseline model during this period. Similarly, the post-EU Referendum period coincides with a fall of 79.9% (Naïve Bayes) and 5.8% (Dictionary) in newspaper sentiment indices. Dictionary indices also show substantial sentiment variation across newspapers during the Global Financial Crisis and post-EU Referendum, with higher standard deviations recorded during these time periods than the baseline model.

A higher frequency of economic news is also found during these two time periods, as is a greater degree of heterogeneity in newspaper consensus. More uncertainty (lower consensus) is found during the global financial crisis, but interestingly we find a higher degree of consensus in the post-EU Referendum period. However, our findings suggest a clear distinction between two publication formats during this period: quality newspapers tend to report news in a more

balanced fashion which incorporates a range of opinions, whereas tabloids tend to take a consistently hard-line and polarised stance on the issue.

The paper is structured as follows: The following section explains the construction of the data set and the time series of the news sentiment indicators. Section Three describes the empirical modelling and discusses the results before Section Four concludes.

2. Construction of the News Sentiment Indicator

2.1. News Consumption through Newspapers

This section describes the thirteen newspapers used in the sample. Four newspapers – The Sun, The Daily & Sunday Mail, The Daily & Sunday Mirror, and The Daily Record – are considered tabloid publications, within which policy and economic news tend to be less prevalent. The Daily & Sunday Mail may also be regarded as a middle-market newspaper as it regularly covers important economic news events. The Evening Standard is a free daily newspaper which covers economic and business news.[4] The remaining newspapers are quality (broadsheet) publications, which typically report on key economic events. Of the thirteen newspapers, tabloid publications command a greater readership: The Daily Mail (26%) and The Sun (22%) are most widely read,[5] followed by The Daily Mirror (11%), Metro (9%),[6] The Times (9%) and The Guardian (9%). The analysis is limited to nationally-distributed newspapers, as the percentage of people who source news through national publications is considerably larger than those reading local and regional equivalents (Ofcom, 2016) and the UK public predominantly consume news in either quality (broadsheet), middle-market and popular (tabloid) newspapers.

[4] As from 2018, former Chancellor of the Exchequer George Osborne has been appointed as the Editor of the Evening Standard.
[5] Based on a National Readership Survey in 2016.
[6] Metro was excluded in this study since the time series was only available after 2008, which is smaller than the sample considered.

Figure A.1 in Appendix A shows a substantial degree of variation in readership – both in print and digital format – for selected newspapers in the United Kingdom. Table A.1 (also in Appendix A) classifies news consumption from national newspapers according to demographic characteristics of the readership. Quality newspapers – such as The Guardian, The Times and The Daily Telegraph command a lower readership in comparison to tabloid newspapers. Age is an essential factor in explaining the differences in readership: compared to readers aged 65+, readership amongst those within the 16-34 category is higher for The Sun, The Guardian and The Financial Times.

Social class and education also play a crucial role in determining readership. For example, readers in lower socioeconomic categories primarily tend to consume news as reported in The Sun, The Daily & Sunday Mirror, and The Daily & Sunday Mail. News readership tends to be more varied amongst those in the higher economic brackets: The Daily & Sunday Mail is the most frequently read newspaper, while tabloid publications such as The Sun and Daily & Sunday Mirror are also popular; albeit to a much lower extent than is the case for the lower economic bracket.[7] It is worth noting that quality newspapers are almost exclusively read by people in the higher economic category.

2.2. Construction of the News Sentiment Indicator

In this section an overview of the dataset is provided and the process of constructing media-based sentiment indicators is elaborated on. First, relevant newspaper articles from UK newspapers containing specific keywords are downloaded from the Lexis Nexis database.[8] Specifically, relevant articles within each of the thirteen newspapers are identified based on an intersection of three

[7] Other important newspapers in this group include The Guardian, The Times, The Sunday Times, and The Daily Telegraph.

[8] The newspapers selected in the dataset (illustrated in Table 1) are collected based on their availability, wide readership and coverage of economic news.

search categories; where the first category (A) refers to the object of interest, the second category (B) relates to policy instruments and institutions and the third category (C) relates to the country that the article refers to. The categories of the search terms are as follows:

A: "Economy" or "Economic"
B: "Policy" or "Taxes" or "Budget" or "Bank of England"
C: "UK" or "Britain" or "British"

The selected articles are intended to depict the state of the UK economy. The words associated with category B are common words one encounters when describing the economic landscape of the country.[9]

Second, a Naïve Bayes machine learning classifier[10] is applied to identify relevant and non-relevant articles in the sample.[11] Before this process takes place, the classifier is first trained using a sample of 600 articles manually classified by the researchers as relevant or non-relevant. Articles classified as non-relevant by the machine learning algorithm are subsequently removed from the dataset. The discussion that follows is based on this filtered dataset.

The relevance of the keywords used to identify relevant articles is subject to a robustness check to ensure that the predefined keywords convey information about the economy. The set of keywords used in categories A and C are not explicitly related to policy, whereas the keywords used in category B restricts the selection of articles to those that are relevant to the theme of policy.[12] One of the potential measurement issues which arises is whether the inclusion of category B restricts

[9] A robustness check explained later shows that by using these specific search terms, the selected articles are appropriate for the model.
[10] An outline of the Naïve Bayes algorithm used in this study is provided later in this paper.
[11] As a result of this additional filtering process, the number of articles used in the final sample is reduced by 29.4%.
[12] It is found that the number of articles increases by 50% when category B is removed from the search, and only search terms specified in categories A & C are specified.

'relevant' articles. Table A.2[13] shows the percentage of 'relevant' articles missed out by including category B, when the Bank of England minutes are released. In a majority of cases the proportion of relevant articles missed is lower than 2%, and thus suggests that measurement errors resulting from keyword selection are marginal.

Table 1. Sample Characteristics

Newswires	Number of articles	Percentage of total sample	Period
Financial Times	133,438	28.85	Sep 1991-Dec 2016
Daily & Sunday Mail	33,238	7.19	Jan 1992-Dec 2016
Daily & Sunday Mirror	9,231	2.00	Jun 1995-Dec 2016
Daily Record	6,278	1.36	Jan 1994-Dec 2016
Daily Telegraph	37,138	8.03	Oct 2000-Dec 2016
Guardian	67,871	14.68	Jan 1990-Dec 2016
Independent	61,459	13.29	Jan 1990-Dec 2016
Sun	8,517	1.84	Jun 2000- Dec 2016
Times	51,319	11.10	Jan 1990-Dec 2016
Observer	8,573	1.85	Oct 1993-Dec 2016
Sunday Telegraph	6,613	1.43	Nov 2000-Dec 2016
Sunday Times	17,020	3.68	Jan 1990-Dec 2016
Evening Standard	21,751	4.70	Jan 1992-Dec 2016
Total	**462,446**	**100.00**	

Notes: This table illustrates the number of articles collected based on the search criteria. It also shows the total percentage of the whole sample, and the time period for which newspaper data was available.

Important economic events – such as the global financial crisis of 2007-2008 – appear to attract newspaper readership. Figure 1 shows the annual frequency of articles relevant to the state of the economy within the three main broadsheet newspapers, while Figure 2 shows similar data for tabloid and middle-market

[13] Table A.2 is located in the accompanying online appendices.

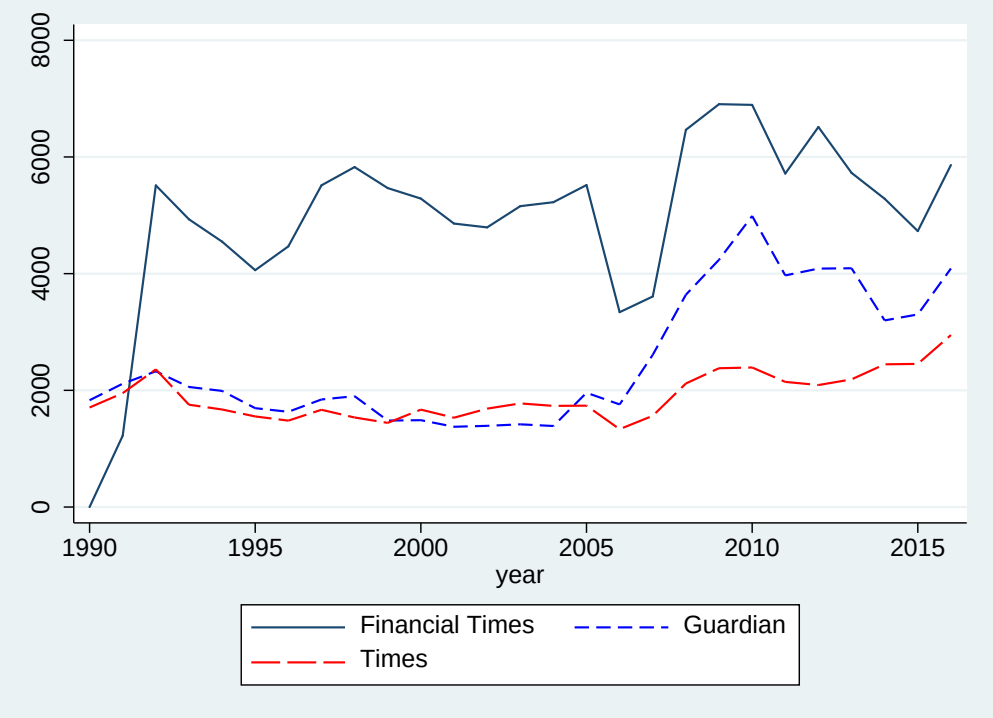

Figure 1. This figure plots the annual number of articles collected using the specified keyword search for the three most popular broadsheet publications.

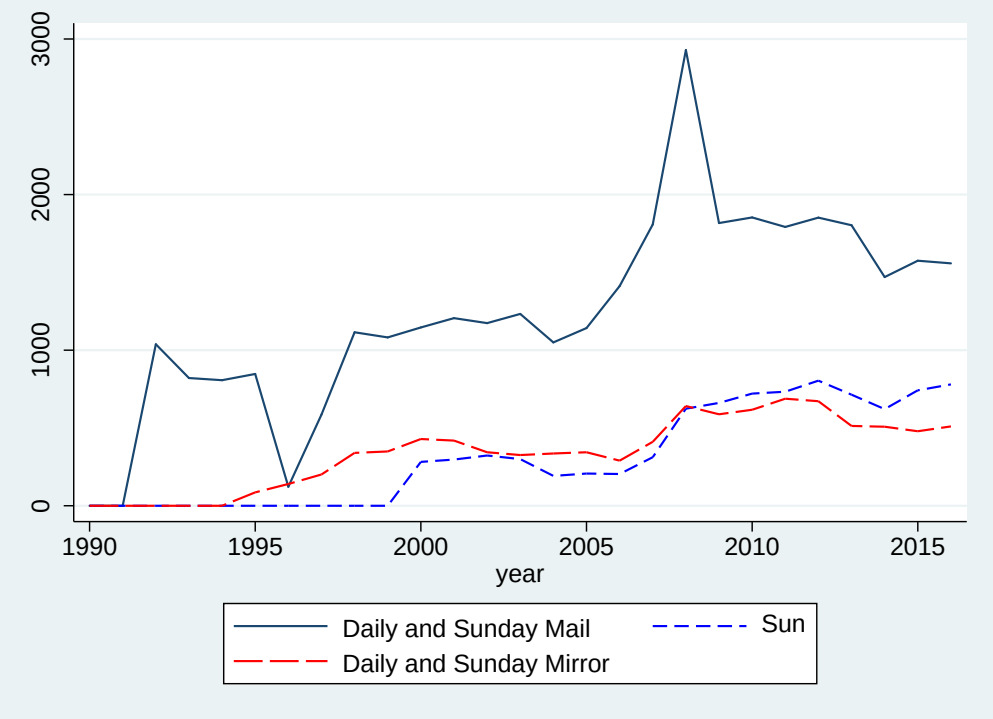

Figure 2. This figure plots the annual number of articles collected using the specified keyword search for the three most popular tabloid and middle-market publications.

newspapers. Prior to the global financial crisis, comparatively fewer articles referencing the economy are observed. It is also evident that events which may impact on future economic conditions – such as the outcome of the EU Referendum – leads to an influx of relevant news articles.

Overall, tabloid publications report economic news less frequently than their middle-market counterparts. The number of relevant articles spikes during the global financial crisis (2007-2009) for both middle-market and tabloid newspapers, although these increases are smaller in magnitude than quality newspapers, and this trend appears to smooth out in the period following the crisis. A particularly high spike in economic news frequency is identified for The Daily & Sunday Mail.

2.2.1. Measuring Economic Sentiment from Newspaper Articles

Studies looking at the influence of news media typically measure sentiment using intensity, searching for terms that relate to positive or negative mood states and counting the frequency with which these words appear in the corpus. The novelty of the approach employed in this paper is that the tonality of newspaper articles is derived using two distinct techniques: Dictionary and Naïve Bayes. The focus on tonality across a broad variety of newspapers is necessary for the current study, given that it deals with broader economic conditions and addresses a reader's perspective on the news article. Also, the influence of news is limited to the readership of the newspapers, which is segmented across the population. Sample characteristics for each newspaper are summarized in Table 1.

For each newspaper, both Naïve Bayes and Dictionary techniques are employed in the construction of sentiment indices. The Naïve Bayes method allows for human subjectivity in the interpretation of newspaper article sentiment, while the Dictionary method assigns a sentiment value based on the pre-defined emotional value of a word.

Firstly, using a Naïve Bayes classifier, articles in the corpus are classified on a three-point scale, where – based on the sentiment conveyed within the content – articles judged as providing positive, neutral and negative outlook are assigned a score of 1, 0 and -1 respectively. Prior to this, the Naïve Bayes sentiment classifier is trained using a subset of articles published in the City AM newspaper,[14] where such articles are manually audited and assigned a score agreed upon by the three authors of this paper. In total, the sentiment training set is comprised of 1,050 articles (350 articles per category) which is consistent with other studies using this method (Tobback et. al. 2016).[15] Once trained, the classifier is then used to classify the corpus of relevant newspaper articles. The Naïve Bayes procedure reflects a "bag of words" approach, wherein words common to a specific sentiment category are identified irrespective of grammar and word order. In doing so, the classifier detects the probability of a new article in the corpus belonging to each of the relevant categories. The theoretical approach to classifying textual data is commonly adopted within computer science literature, and has recently been applied in economics (Gentzkow et. al., 2017; Bholat et. al., 2015).

The accuracy of the Naïve Bayes classifier is evaluated using the confusion matrix technique detailed by Kohavi and Provost (1998). Using an out-of-sample validation set, the Naïve Bayes classifier achieves an accuracy of 72.0%,[16] a figure broadly in keeping with prior research. For example, Wang et al. (2014) achieve an

[14] City AM typically reports economic and financial news on a daily basis and thus provides phrasing and wording which is commonly used to communicate economic and financial news in other newspapers.

[15] Note that the Naïve Bayes classifier referred to here is used to assign sentiment scores to news article using a training set of 1,050 articles classified as conveying positive, negative, or neutral sentiment. This classifier therefore differs from the Naïve Bayes classifier referred to earlier in Section 2.2, which was used to distinguish relevant and non-relevant articles using a training set of 600 articles which were manually classified as relevant or non-relevant.

[16] Accuracy is calculated as the proportion of true positives (e.g. articles correctly identified by the Naïve Bayes classifier as positive, neutral or negative). Similar levels of precision (exactness) and recall (sensitivity) measures (70.35% and 75.50%, respectively) are identified. Further details of these measures can be found in Kohavi and Provost (1998).

accuracy of 76.2% using a corpus of short messages on a financial social media platform, whereas Koppel and Shtrimberg (2006) achieve 65.9%.[17]

As a second measure for measuring tonality within newspaper articles, Loughran and Macdonald's (2011) compiled dictionary of terms is also considered.[18] Words within the dictionary are accompanied by a score based on whether they are positive (+1) or negative (-1). In this setting, the Dictionary technique will measure the outcome of an article based on the number of words with positive and negative associations. The sentiment 'score' assigned to a given article is computed as:

$$S = \frac{N^+ - N^-}{N^+ + N^- + N^0}, \tag{1}$$

where S represents the assigned article sentiment, derived using the number of words in the article associated with positive (N^+), negative (N^-) and neutral (N^0) sentiment. This approach is ideal due to linguistic differences in the communication of economic news. Indeed, 73.90% out-of-sample accuracy is detected using this technique; marginally higher than the ability of the Naïve Bayes classifier.[19] Once all articles are classified according to both the Naïve Bayes and Dictionary algorithms, an aggregated sentiment score for each newspaper is then computed by averaging the sentiment scores of individual articles at monthly intervals.

[17] A number of studies using similar techniques report higher accuracy using an in-sample validation set (Sprenger et al., 2014). For example, Koppel and Shtrimberg (2006) achieve 70.3% accuracy in-sample and 65.9% accuracy out-of-sample.

[18] The Loughran and Macdonald (2011) dictionary is used for textual analysis for finance and business topics, and the key terms used may therefore be closer to words in an economic context.

[19] However, the measures of precision (69.45%) and recall (63.11%) are lower than that of the Naïve Bayes classifier; namely due to the dictionary algorithm's decreased ability in identifying positive articles, and the high proportion of negative and neutral articles within the random selection of articles forming the validation set.

2.2.2. Descriptive Statistics of Newspaper Sentiment Indices

Table 2 illustrates the summary statistics from each newspaper based on monthly arithmetic mean aggregation. Inspection of the table suggests that the Naïve Bayes technique produces a positive mean for The Financial Times, whereas all other newspapers have negative means using this method. Negative means are particularly strong for tabloid newspapers – such as The Sun, and The Daily & Sunday Mirror – which are characterised by large negative means. As we will discuss later in this paper, this may be due to the fact that tabloids are generally more likely to report on economic outlook during periods of uncertainty or negativity. When the Naïve Bayes method is substituted with Loughran and Macdonald's (2011) Dictionary method, sentiment scores for all newspapers are negative.

Time series of the sentiment indices for each newspaper using both the Naïve Bayes and Dictionary methods are provided in Figures B.1 and B.2, respectively.[20] Significant volatility in the indices is noticed; especially around political and economic events such as Labour's election (1998), the global financial crisis (2007-2008), and the Greek debt crisis (2010). The Naïve Bayes time series suggest that sentiment scores tend to be consistently lower for tabloids, in comparison to quality publications.

Appendix B reports the correlation of sentiment indices across newspapers. The correlations tend to be positive but differ in nature dependent on the textual analysis method used. The Dictionary method (Table B.2) method produces strong positive correlations (a correlation coefficient around 0.700 and above) between newspapers such as The Daily Telegraph, The Evening Standard, The Financial Times, The Guardian, The Independent, Daily & Sunday Mail and The Times. All of these sources – with the exception of Daily & Sunday Mail – are defined as broadsheet

[20] Due to the large number of individual time series, Figures B.1 and B.2 are found in the accompanying online appendices.

Table 2: Summary Statistics of Newspaper Sentiment Indicators

	T	Mean	Std Dev	5%	95%
Panel A: Naïve Bayes					
Daily & Sunday Mail	300	-0.1242	0.1828	-0.3898	0.1642
Daily & Sunday Mirror	259	-0.2953	0.2023	-0.6081	0.0858
Daily Record	276	-0.2060	0.2619	-0.5573	0.2143
Daily Telegraph	195	-0.0977	0.1151	-0.2836	0.0985
Evening Standard	300	-0.0118	0.1883	-0.3284	0.2732
Financial Times	304	0.0694	0.1054	-0.1008	0.2507
Guardian	324	-0.1155	0.1048	-0.2947	0.0543
Independent	324	-0.0465	0.1355	-0.2759	0.1970
Observer	279	-0.1658	0.1705	-0.4324	0.1058
Sun	199	-0.3832	0.1807	-0.6673	-0.0923
Sunday Telegraph	194	-0.0866	0.1873	-0.3846	0.1919
Sunday Times	324	-0.0944	0.1497	-0.3130	0.1563
Times	324	-0.0317	0.1165	-0.2160	0.1722
Panel B: Dictionary					
Daily & Sunday Mail	300	-1.5939	0.3697	-2.2050	-1.0373
Daily & Sunday Mirror	259	-1.5778	0.6234	-2.6123	-0.5503
Daily Record	276	-1.2166	0.6349	-2.1375	-0.1718
Daily Telegraph	195	-1.5665	0.3428	-2.2406	-1.0749
Evening Standard	300	-1.5786	0.4213	-2.3289	-0.9004
Financial Times	304	-1.5829	0.3032	-2.1062	-1.1534
Guardian	324	-1.6529	0.2987	-2.1906	-1.2008
Independent	324	-1.5287	0.3167	-2.0656	-1.0495
Observer	279	-1.5045	0.4316	-2.2557	-0.8602
Sun	199	-1.6006	0.5186	-2.4575	-0.7739
Sunday Telegraph	194	-1.3364	0.4060	-2.0292	-0.7418
Sunday Times	324	-1.2388	0.3708	-1.8475	-0.6476
Times	324	-1.4823	0.3057	-2.0016	-1.0418

Notes: The table illustrates the descriptive statistics for each of the newspapers based on the Naïve Bayes (Panel A) and Dictionary (Panel B) classification. T represents the length of each time series. The summary statistics considered are the mean, standard deviation, and 5% and 95% percentiles.

newspapers. Tabloids, such as The Sun, Daily & Sunday Mirror and The Daily Record are less correlated with other newspapers and between each other, with most of the correlation coefficients not exceeding 0.500. This suggests that quality newspapers portray economic conditions more consistently than tabloids. In the case of Naïve Bayes (Table B.1), the results show a positive correlation across many newspapers.

Low correlations with other newspapers (which is the case for tabloids mainly) may be due to newspapers weighing on the relevance of articles which appeal to their readership. The choice of economic news presented in such newspapers reflects the relevance to the demands of their readers (Gentzkow and Shapiro, 2010).[21] There is evidence of moderately high positive correlation between Naïve Bayes and Dictionary sentiment scores for the full sample, with coefficients ranging from 0.188 (The Sun) to 0.604 (The Guardian).[22] This is encouraging given that both techniques employ different approaches: Naïve Bayes is a supervised learning technique whereas Dictionary is a frequency count of specific words with positive or negative connotation. Even though the techniques entirely differ in the way they consider text and emphasize its different features, they still agree to a certain extent in their measurement of the sentiment. This provides an argument about the validity of the analytical approach used in the sentiment classification.

3. Results and Discussion

In this section, a series of hypotheses using the newly constructed indices are tested. The first hypothesis tests for evidence of newspaper-specific effects in the

[21] For instance, some newspapers may report news more relevant to households (such as dynamics of inflation or wages) as opposed to other newspapers more focussed on matters related to industry or the financial sector (such as GDP, investment, or trade).

[22] A full list of correlation coefficients are shown in Table C.3, which can be found in the accompanying online appendix.

sentiment series across the whole sample. Secondly, we test for heterogeneity amongst newspapers in their reporting of economic news during the periods (i) covering the global financial crisis and (ii) following the UK's EU Referendum, in terms of mean sentiment levels, volume and consensus. Thirdly, we test whether shifts in mean sentiment scores are detected across important economic events. Fourthly, we investigate whether there is evidence of movements in consumer sentiment indices which can be traced from newspapers. Finally, we consider the relevance of the newspaper sentiment scores for forecasting.

3.1 Newspaper fixed effects

Heterogeneity across newspapers after accounting for time effects can be investigated using a similar model to the fixed effects model of Shapiro et al. (2017), where newspaper article scores are regressed against time dummies and newspaper effects of the form:

$$S_{i,t,j} = \beta_{1,t} TIME_t + \beta_{2,j} NP_j + \varepsilon_{i,t,j}, \qquad (2)$$

where $S_{i,t,j}$ is the sentiment score assigned to article i, from newspaper j at time t. $TIME_t$ is a time dummy which equals to one if the article comes from month t, and NP_j is the newspaper j dummy. The estimates of $\beta_{2,j}$ show the mean article sentiment score from newspaper j conditioning on time effects. $\beta_{1,t}$ can also be interpreted as a combined monthly sentiment score in month t after filtering out newspaper effects.

Results of the estimates are provided in Table 3. The table shows that all parameters are statistically significant for the Dictionary method. In contrast, both statistically significant and non-significant parameters are present for the Naïve Bayes method. The newspapers with non-significant parameters are quality newspapers, such as The Telegraph, The Guardian, The Independent and The Times, Evening Standard, Sunday Telegraph, Sunday Times. An insignificant

newspaper specific parameter indicates that there is little deviation from the combined newspaper effects (time dummy). Thus, quality newspapers tend to be more objective in the sense that the implied sentiment in their reporting of economic news is close to the average for all the newspapers. It is worth noting a statistically significant effect for The Financial Times, which is a newspaper with particularly rich coverage of economic news.

Table 3. Newspaper Fixed Effects Parameters

	Naive Bayes (N=462,446)			**Dictionary** (N=462,446)		
Newspaper	**Est.**	**Std Error**	**P-val**	**Est.**	**Std Error**	**P-val**
Daily & Sunday Mail	-0.114	0.054	0.034	-1.534	0.093	0.000
Daily & Sunday Mirror	-0.256	0.054	0.000	-1.564	0.096	0.000
Daily Record	-0.183	0.057	0.001	-1.221	0.097	0.000
Daily Telegraph	-0.030	0.053	0.577	-1.422	0.093	0.000
Evening Standard	0.019	0.054	0.721	-1.520	0.094	0.000
Financial Times	0.112	0.053	0.035	-1.511	0.093	0.000
Guardian	-0.066	0.053	0.213	-1.552	0.093	0.000
Independent	-0.021	0.053	0.691	-1.469	0.093	0.000
Observer	-0.107	0.054	0.048	-1.449	0.095	0.000
Sun	-0.323	0.054	0.000	-1.408	0.096	0.000
Sunday Telegraph	-0.008	0.055	0.886	-1.155	0.095	0.000
Sunday Times	-0.049	0.053	0.355	-1.133	0.094	0.000
Times	0.008	0.053	0.879	-1.395	0.093	0.000

Notes: The table illustrates the fixed effects parameters estimated from Equation (2) for Naïve Bayes and Dictionary sentiment scores. The first column illustrates the newspapers. The columns of each subcategory, in turn, illustrate the estimated parameter ('Est.'), (robust) standard error ('Std Error') and p-value ('P-val').

There is little agreement with respect to the categories of tabloids and quality newspapers when the two techniques are compared. The Naïve Bayes classifier shows a strong negative average for tabloids (The Sun, The Daily and Sunday Mirror and The Daily Record) of -0.254 against an average of -0.021 for quality

newspapers[23]. The Dictionary classifier shows only a marginal difference between both.

While the estimates derived using Equation (2) show heterogeneity across newspapers and time periods, they do not necessarily indicate that the differences from each other are statistically significant. To address this issue a test of joint significance (with the null hypothesis being that the parameters from any two newspapers being equal) is performed, with results reported in Appendix C. For the Naïve Bayes model (Table C.1), it is found that The Observer and The Daily & Sunday Mail do not differ significantly (p-value of 0.511) in terms of newspaper-specific effects noting that these two newspapers have similar sample sizes. Strong non-rejection levels are also noted amongst quality newspapers.[24] The Dictionary method offers additional insightful findings. For example, there is no evidence of significantly different effects between The Daily & Sunday Mail and The Evening Standard (p-value = 0.351), The Guardian and The Daily & Sunday Mirror (0.511), The Daily Telegraph and The Sun (0.480), The Evening Standard and The Financial Times (0.477), and The Sunday Telegraph and The Sunday Times (0.369).

3.2 Sentiment Scores and External Events

This section examines the extent of variation in newspaper sentiment indices during two specific economic events: the global financial crisis of 2007-2008 and the UK's EU Referendum vote (including the period thereafter). These events are distinct in that the first concerns a period of global economic crisis while the second

[23] In this definition, The Financial Times, The Observer and Daily & Sunday Mail have been excluded.

[24] A similar lack of difference between quality publications is identified in the case of The Evening Standard and The Sunday Telegraph (0.197), The Independent and The Sunday Telegraph (0.214), and The Sunday Telegraph and The Times (0.142). In these instances, the differences are not found to be statistically significant.

pertains to a period of future economic and political uncertainty caused by domestic factors.

We test whether sentiment scores are significantly different across newspapers by estimating a similar model in which article sentiment scores are regressed against newspaper dummies, financial crisis and post-EU Referendum dummies, and interaction terms. The model considered is an extension of Equation (2):

$$S_{i,t,j} = \beta_1 FC + \beta_2 ER + \beta_{3,j} NP_j + \beta_{4,j} FC \cdot NP_j + \beta_{5,j} ER \cdot NP_j + \varepsilon_{i,t,j}, \qquad (3)$$

where *FC* is a dummy variable representing the financial crisis period, *ER* is a dummy variable accounting for the six months after the EU Referendum vote (June 2016-December 2016), *FC·NP* is an interaction dummy of each newspaper with the financial crisis period, and *ER·NP* is an interaction dummy of each newspaper with the post-Referendum dates.

Table 4 reports significant newspaper effects. After accounting for exogenous events, economic news reporting (or the sentiment inferred from news articles) is significantly different across newspapers using both textual techniques; a finding which is expected given earlier analysis. The global financial crisis (β_1) parameters are more negative than post-EU Referendum (β_2) for both techniques. The mean values across all newspapers in the baseline model ($\beta_{3,j}$) are -0.126 (Naïve Bayes) and -1.47 (Dictionary), with standard deviations of 0.126 (Naïve Bayes) and 0.130 (Dictionary).

For the global financial crisis and the post-EU Referendum period, the mean sentiment for the Naïve Bayes is substantially larger at 0.129 and 0.117 respectively, in comparison to Dictionary values of -1.833 and -1.529. There is a substantial increase in standard deviation across newspapers for these events using Dictionary.

Table 4. Estimation of Newspaper Responses to Events

	Naïve Bayes (N=462,446)			**Dictionary** (N=462,446)		
Global Financial Crisis (β_1)	-0.212***			-0.384***		
Post-EU Referendum (β_2)	-0.149***			-0.200***		
	$\boldsymbol{\beta_{3,j}}$	$\boldsymbol{\beta_{4,j}}$	$\boldsymbol{\beta_{5,j}}$	$\boldsymbol{\beta_{3,j}}$	$\boldsymbol{\beta_{4,j}}$	$\boldsymbol{\beta_{5,j}}$
Daily & Sunday Mail	-0.125**	-	-	-1.581***	-	-
Daily & Sunday Mirror	-0.307***	0.113***	0.024	-1.617***	0.014	-0.279**
Daily Record	-0.258***	0.225***	0.016	-1.322***	0.221***	-0.125
Daily Telegraph	-0.097***	0.139***	0.111***	-1.514***	-0.010	0.343***
Evening Standard	0.012**	0.013	-0.002	-1.507***	-0.234***	0.138
Financial Times	0.063***	0.214***	0.061**	-1.525***	-0.102***	0.164**
Guardian	-0.118***	0.153***	0.072**	-1.642***	0.128***	0.229***
Independent	-0.069***	0.193***	-0.046	-1.549***	0.187***	-0.071
Observer	-0.129**	0.055***	-0.006	-1.425***	-0.216***	0.015
Sun	-0.420***	0.203***	0.204***	-1.541***	0.096**	0.347***
Sunday Telegraph	-0.072***	0.141***	0.053	-1.240***	0.035	0.086
Sunday Times	-0.095***	0.146***	0.042	-1.221***	0.162***	0.104
Times	-0.029***	0.118***	0.094***	-1.445***	0.000	0.232***

Notes: ***, ** and * denote statistical significance at the 1%, 5% and 10% level respectively. The table illustrates the estimated coefficients from Equation (3) (with robust standard errors). The top two columns illustrate the coefficients for the intercept term (The Daily Mail as the baseline), the dummy for the Global Financial Crisis years (β_1), and the post-EU Referendum dummy (β_2) for both textual techniques. The second part of the table illustrates the coefficients from the newspaper fixed effects$(\beta_{3,j})$, and interactive dummies for the Global Financial Crisis and post-EU Referendum.

To better understand the extent of heterogeneity across newspapers and events, the overall newspaper-event effect measure is computed based on the parameters from Table 4. These effects are illustrated in Figures 3 and 4. Newspaper-specific coefficients during the global financial crisis period, the post-EU Referendum period and the baseline sample are shown for the Naïve Bayes (Figure 3) and Dictionary (Figure 4) indices respectively. In the case of both sentiment classifiers, negative sentiment is more prevalent during both the global financial crisis except for The Financial Times. Notable exceptions to this are The Daily Record (for Naïve Bayes) and The Sun (Naïve Bayes and Dictionary), which are both tabloid

publications, and – perhaps more surprisingly – The Financial Times (Naïve Bayes and Dictionary), The Guardian (Dictionary) and The Daily Telegraph (Dictionary).

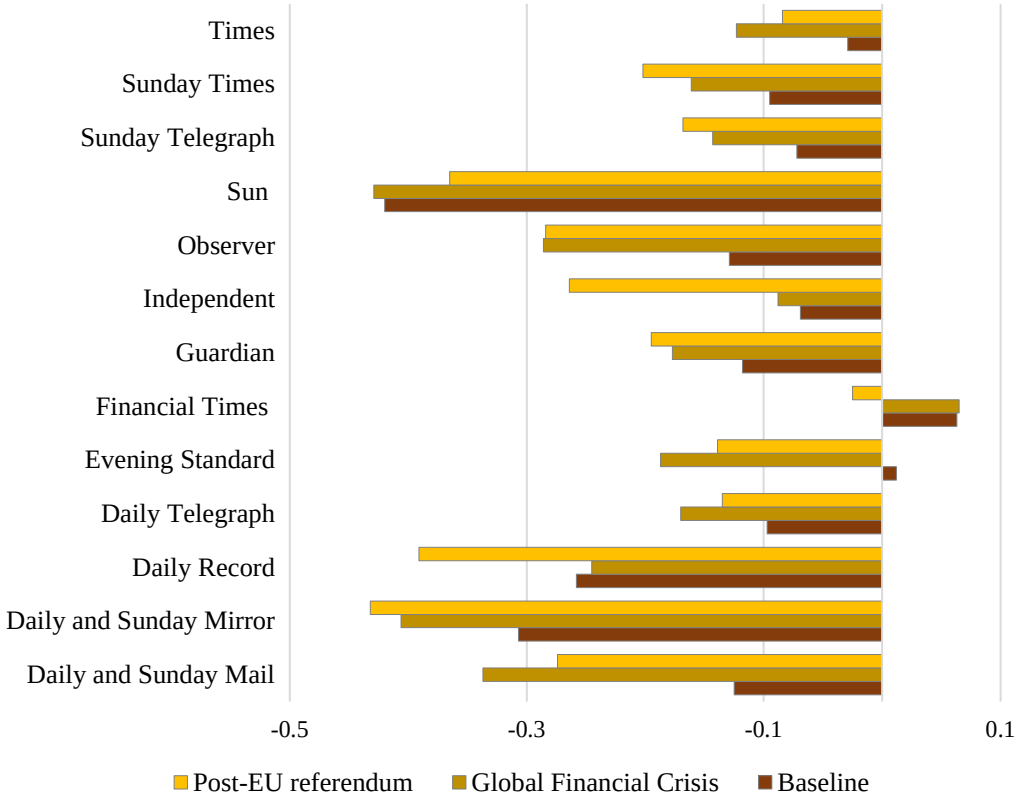

Figure 3. The figure shows the computed coefficient, based on estimates in equation (3), using the Naïve Bayes technique for (i) the financial crisis period, (ii) the post-EU Referendum vote period, and (iii) the baseline model.

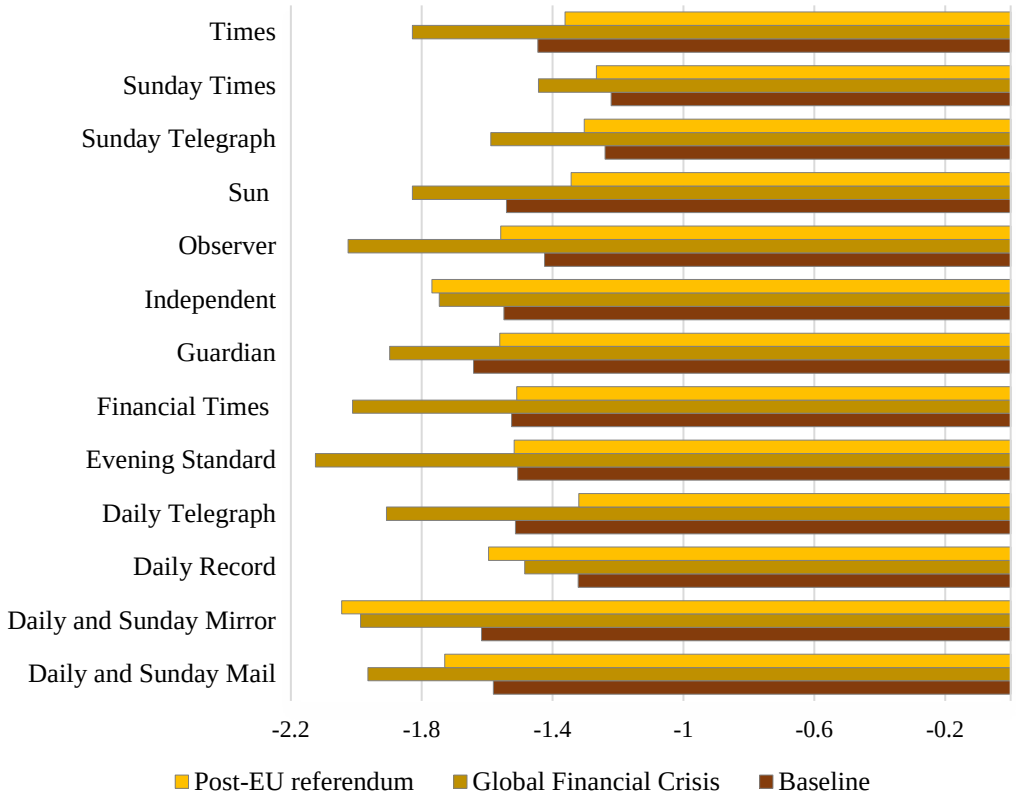

Figure 4. The figure shows the computed coefficient, based on estimates in equation (3), using the Dictionary technique for (i) the financial crisis period, (ii) the post-EU Referendum vote period, and (iii) the baseline model.

The Naïve Bayes and Dictionary methods show similarities in terms of sizeable importance of the financial crisis and the post-EU Referendum period, yet there also exist some notable differences. Consensus between the textual techniques (where both suggest stronger negative tonality for the global financial crisis than the post-EU Referendum period) is found to occur in the case of The Times, The Sun, The Evening Standard and The Daily Telegraph, The Observer and The Daily and Sunday Mail. Interestingly, quality newspapers such as The Sunday Telegraph, The Sunday Times, The Guardian and The Financial Times demonstrate stronger negative tonality during the global financial crisis using the Dictionary method, but not when using Naïve Bayes. In other cases, for both textual techniques the post-EU Referendum period appears to be portrayed with more negative sentiment than

the global financial crisis. This applies to The Independent, The Daily Record, and The Daily & Sunday Mirror.

3.3 Heterogeneity in Volume

We also consider whether there are significant effects in terms of the frequency of relevant news articles published during the global financial crisis and post-EU Referendum periods. To that end, we regress the number of articles per newspaper each month against the newspaper dummy, the global financial crisis and post-EU Referendum dummies, and interaction dummies with both events:

$$Count_{t,j} = \beta_1 FC + \beta_2 ER + \beta_{3,j} NP_j + \beta_{4,j} FC \cdot NP_j + \beta_{5,j} ER \cdot NP_j + \varepsilon_{t,j}, \qquad (4)$$

where $Count_{t,j}$ is the number of articles in newspaper j during month t. The hypothesis is that around certain events of turmoil, the volume of articles may increase in order to reiterate the importance of such events.

Table 5. Volume Effects

	Count (N=324)		
Global Financial Crisis (β_1)	82.908***		
Post-EU Referendum (β_2)	31.075***		
	$\beta_{3,i}$	$\beta_{4,i}$	$\beta_{5,i}$
Daily & Sunday Mail	99.07***		
Daily & Sunday Mirror	32.74***	-66.58***	-19.53
Daily Record	19.91***	-65.42***	-15.84
Daily Telegraph	175.73***	-18.07	-0.66
Evening Standard	66.49***	-40.35*	-22.71
Financial Times	419.19***	38.71	114.45***
Guardian	187.83***	71.41*	67.24**
Independent	188.30***	-115.08***	221.77***
Observer	29.59***	-73.57***	-40.53
Sun	39.20***	-70.01***	-4.27
Sunday Telegraph	31.00***	-70.15***	-20.22
Sunday Times	47.19***	-45.95**	-0.26
Times	151.54***	-49.27**	89.25***

Notes: The table illustrates the coefficients from estimating Equation (4) using robust standard errors. The parameters illustrate the number of articles containing the keywords specified in Section 2.2 across the newspapers.

The results show that news article volume effects are statistically significant for all newspapers (parameter β_3), where newspapers differ in terms of the number of articles they publish per month. Tabloids have the lowest volume of economic articles followed by Sunday prints. The difference in economic articles published during the baseline period and the global financial crisis is the sum of $(\beta_1 + \beta_4)$.[25] During the global financial crisis, the newspapers tend to publish substantially more articles reporting economic news than in other times, as depicted by a highly statistically significant and positive parameter β_2 (82.908). By and large, The Guardian and The Financial Times show a higher number of published articles during the global financial crisis (on top of the newspaper-specific effects), suggesting increased economic activity reporting at the time. In this context, The Independent is particularly interesting as it has a highly negative coefficient during the financial crisis period.[26]

Regarding the period following the UK's EU Referendum, the model indicates that there are substantially more articles published in quality newspapers such as The Financial Times, The Guardian, The Independent and The Times. After computing the average changes across the different categories of newspapers, the global financial crisis witnesses an increase in article volume of 50.9% for tabloids and 36.9% for quality newspapers. However, the post-EU Referendum shows an increase in volume of 58.0% and 65.1%, respectively. This finding shows heterogeneity in response to two different types of events as far as uncertainty is concerned. Uncertainty regarding future economic outcomes following the EU Referendum may be stronger, which quality newspapers are more likely to address in their communications.

[25] For the post-EU Referendum, the effect is computed as $(\beta_2 + \beta_5)$.
[26] During the 2007-2008 period, The Independent underwent various changes in their organisational structure, including staff cuts and a merger, which perhaps may explain why comparatively it is found that the coefficient is negative.

3.4 Consensus and mixed signals

The standard deviation of sentiment scores in each month can be considered as a measure of consensus across articles, or the degree of agreement to which the newspaper has arrived regarding the current economic state. Newspapers can offer mixed opinions on current economic conditions, current economic policies or anticipation for the future. In the current study, consensus is simply defined using a standard deviation of sentiment scores, with low values interpreted as reflecting a high degree of consensus. During times of low ambiguity, it can be expected that the standard deviation across sentiment scores will be comparatively lower, indicating a higher degree of consensus. On the other hand, a high standard deviation is indicative of mixed signals and a lack of consensus about current or future economic conditions. In this context it is interesting to establish the level of consensus amongst newspapers during times of economic uncertainty, such as the global financial crisis or the post-EU Referendum period. The equation put forward to investigate this problem is given by:

$$SD_{j,t} = \beta_1 FC + \beta_2 ER + \beta_{3,j} NP_j + \beta_{4,j} FC \cdot NP_j + +\beta_{5,j} ER \cdot NP_j + \varepsilon_{j,t}, \quad (5)$$

where $SD_{j,t}$ is the standard deviation across articles for newspaper *j* and for month *t*. The estimation results are presented in Table 6.

Table 6. Standard Deviation Effects

	Naïve Bayes (N=462,446)			**Dictionary** (N=462,446)		
Global Financial Crisis(β_1)		-0.010			0.115***	
post-EU Referendum(β_2)		-0.088**			0.104***	
	$\beta_{3,i}$	$\beta_{4,i}$	$\beta_{5,i}$	$\beta_{3,i}$	$\beta_{4,i}$	$\beta_{5,i}$
Daily & Sunday Mail	0.819***	-	-	1.655***	-	-
Daily & Sunday Mirror	0.804***	-0.021	0.053	1.958***	-0.084*	-0.176***
Daily Record	0.810***	0.066***	-0.032	1.769***	0.156**	-0.255***
Daily Telegraph	0.847***	0.035**	0.071*	1.734***	-0.038	-0.305***
Evening Standard	0.863***	0.009	0.080**	1.704***	-0.040	-0.136***
Financial Times	0.857***	0.030*	0.095**	1.721***	-0.125**	-0.312***
Guardian	0.768***	0.027	0.011	1.602***	0.018	-0.238***
Independent	0.813***	0.023	-0.017	1.596***	-0.130***	-0.028
Observer	0.708***	0.013	-0.071*	1.468***	0.059	-0.447***
Sun	0.770***	0.026	0.119***	2.028***	-0.179***	-0.277***
Sunday Telegraph	0.773***	0.020	0.120***	1.463***	-0.071	-0.146***
Sunday Times	0.775***	0.020	0.039	1.544***	-0.078*	-0.090**
Times	0.829***	0.000	0.088**	1.659***	-0.115***	-0.178**

Notes: The results show the estimates from equation (5) for each newspaper using robust standard errors. The dependent variable is the degree consensus reached on economic matters which is proxied by the standard deviation across articles over the month. The estimates explain whether there are significant newspaper fixed effects across the different events examined for both techniques.

At a baseline level, the model shows statistically significant newspaper-specific effects ($\beta_{3,i}$) for all newspapers using both textual techniques. It also indicates statistically significant positive parameters for the global financial crisis and the post-EU Referendum period using the Dictionary indices, and a statistically significant negative parameter for the post-EU Referendum period using Naïve Bayes indices. This means that the Dictionary technique detects a lower degree of consensus among newspapers during the specific events considered. In turn, the Naïve Bayes measure suggests an increase in consensus for the post-EU Referendum period only.

Analysis of the newspaper interaction terms with the specific events brings additional insights. For both the Dictionary and Naïve Bayes indices, fewer newspapers interact with the events in a statistically significant way as compared

to the newspaper-only fixed effects. To facilitate interpretation, the overall newspaper-specific effects are computed for both events, by aggregating the parameters reported in Table 6 for the global financial crisis and the post-EU Referendum period (Table 7).

Table 7 shows consensus scores for Naïve Bayes and Dictionary indices over three periods: (i) the baseline case, which covers the entire window without the effects of the specific events; (ii) for the global financial crisis period; and (iii) for the post-EU Referendum period. The results are interesting, in the sense that both Naïve Bayes and Dictionary show an increase in computed parameters (hence, less consensus) during the financial crisis. During the global crisis, newspapers exhibit uncertainty and ambiguity which can be seen in higher means for both Naïve Bayes (0.812) and Dictionary (1.752). However, the opposite is true during post-EU Referendum period, where the mean consensus is better than in the baseline model. Moreover, the standard deviation of the means across newspapers is higher. This result reflects the nature of the sample, where some newspapers may take a polarised position concerning the EU Referendum results ('Remain' versus 'Leave'), while others may simply be more unbiased in reporting the events, which would explain the low means and high standard deviation. The Observer and The Guardian, for instance, may take a political stance on the issue which makes consensus across the articles stronger.

Newspapers such as The Financial Times (0.857), The Evening Standard (0.847), The Daily Telegraph (0.822), The Times (0.821) demonstrate a balanced view of the EU Referendum. Similar findings are observed with the Dictionary - the mean consensus is higher, but with a higher standard deviation. It is also worth recognising that the high standard deviation of the means could also be due to the mixed signals from good economic data in the post-EU Referendum period.

Table 7. Degree of Consensus

	Naïve Bayes			Dictionary		
	Baseline	Financial Crisis	Post-EU Ref	Baseline	Financial Crisis	Post-EU Ref
Daily & Sunday Mail	0.819	0.809	0.723	1.655	1.780	1.759
Daily & Sunday Mirror	0.804	0.773	0.761	1.958	1.989	1.886
Daily Record	0.810	0.866	0.682	1.769	2.04	1.618
Daily Telegraph	0.847	0.872	0.822	1.734	1.811	1.533
Evening Standard	0.863	0.862	0.847	1.704	1.779	1.672
Financial Times	0.857	0.877	0.857	1.721	1.711	1.513
Guardian	0.768	0.785	0.683	1.602	1.735	1.468
Independent	0.813	0.826	0.700	1.596	1.581	1.672
Observer	0.708	0.711	0.541	1.468	1.642	1.125
Sun	0.770	0.786	0.793	2.028	1.964	1.855
Sunday Telegraph	0.773	0.783	0.797	1.463	1.507	1.421
Sunday Times	0.775	0.785	0.718	1.544	1.581	1.558
Times	0.829	0.819	0.821	1.659	1.659	1.585
Mean	0.803	0.812	0.757	1.685	1.752	1.589
Std Deviation	0.042	0.047	0.084	0.161	0.159	0.190

Notes: The table shows the (mean) standard deviation of article sentiment scores per newspaper during three time periods: (i) the full observation window excluding the Financial Crisis period and post-EU Referendum (baseline case); (ii) the period spanning the global financial crisis; and (iii) the period following the UK's EU Referendum.

Table 8 categorizes the newspapers based on the parameters from Table 7, by forming respective similarity clusters. The newspapers with the lowest (highest) standard deviation, between the 1st and 30th percentiles (between the 71st and 100th percentiles) are newspapers with high (low) levels of consensus. The remaining newspapers (between the 31st and 70th percentiles) constitute the 'moderate' consensus category.

The results presented in Table 8 shows that in general the clustering is dependent on the textual technique used. For example, The Times features mainly in the moderate consensus category for the Dictionary method, and mainly in the low category for the Naïve Bayes.

The Naïve Bayes results show that quality newspapers (except for the Sunday prints that are in the high consensus category) are mostly in the low consensus category, with tabloids occupying the high and moderate categories. The Evening Standard, The Financial Times and The Daily Telegraph consistently represent a low degree of consensus over the three different periods. In contrast, The Observer is found to be in the high consensus category during the three periods, with The Guardian and The Sunday Times in two cases out of the three (they are instead found in the moderate consensus category during the global financial crisis and post-EU Referendum, respectively).

The Dictionary method provides a small number of different results from Naïve Bayes, in that tabloid publications (such as The Daily & Sunday Mirror, The Sun, and The Daily Record) are mainly found in the low consensus category, while quality newspapers (such as The Evening Standard, The Daily Telegraph and The Times) are more frequently located in the moderate consensus category. The Observer and The Sunday Telegraph are characterized by a high degree of consensus over all three periods, while The Independent and The Sunday Times are in the high category in two cases out the three (they are both instead found to be in the moderate category during the post-EU Referendum period).

Table 8. Newspaper Clusters on Consensus

	High Consensus *(1-30% percentile)*	**Moderate Consensus** *(31-70% percentile)*	**Low Consensus** *(71-100% percentile)*
		Panel A. Naïve Bayes	
Baseline	Sun Sunday Times Guardian Observer	Daily & Sunday Mail Daily Record Daily & Sunday Mirror Independent Sunday Telegraph	Evening Standard Financial Times Daily Telegraph Times
Global Financial Crisis	Sunday Times Sunday Telegraph Daily & Sunday Mirror Observer	Independent Times Daily & Sunday Mail Sun Guardian	Financial Times Daily Telegraph Daily Record Evening Standard
Post-EU Referendum	Independent Guardian Daily Record Observer	Sunday Telegraph Sun Daily & Sunday Mirror Daily & Sunday Mail Sunday Times	Financial Times Evening Standard Daily Telegraph Times
		Panel B. Dictionary	
Baseline	Sunday Telegraph Observer Sunday Times Independent	Guardian Daily & Sunday Mail Times Evening Standard Daily Telegraph	Financial Times Daily Record Daily & Sunday Mirror Sun
Global Financial Crisis	Sunday Telegraph Independent Sunday Times Observer	Times Financial Times Guardian Daily & Sunday Mail Evening Standard	Daily Telegraph Sun Daily & Sunday Mirror Daily Record
Post-EU Referendum	Observer Sunday Telegraph Guardian Financial Times	Daily Telegraph Sunday Times Times Daily Record Independent	Evening Standard Daily & Sunday Mail Sun Daily & Sunday Mirror

Notes: The table shows clusters of newspapers ranked on percentiles for (i) the baseline case, (ii) the global financial crisis period, (iii) the post-EU Referendum period across the two different textual techniques. The 'high consensus' column includes newspapers where the standard deviation was in the lowest 30% of the sample newspapers (newspapers with high degree of consensus). The 'moderate consensus' column represents the cluster where newspapers had a standard deviation within the 31st to 70th percentile. The 'low consensus' category represents newspapers with a standard deviation within the 71st and 100th percentile and thus reflect a low degree of consensus.

In this context, it is worth noting the relatively low levels of consensus detected in The Financial Times (except for the financial crisis and post-EU Referendum for Dictionary), a quality newspaper specialising in financial and economic topics, and publishing a high number of relevant articles published is high (as evidenced by its sizeable sample in our dataset). This observation is much easier to understand if we assume that, as a quality and reputable newspaper, it publishes many articles that present different (and perhaps sometimes conflicting) views and opinions on the economic situation.

Upon further examination, we find that the increase in consensus from both techniques can be attributed to both quality and tabloids. However, they differ in terms of magnitude. The Naïve Bayes classifier shows, when compared to the baseline parameters, an increase in consensus of 6.2% and 4.9% for tabloids and quality newspapers respectively. This finding is similar for the Dictionary classifier with respective percentages of 6.9% and 3.5%. As mentioned earlier, one potential reason for this finding, is the mixed signal coming from high uncertainty in the future and also the good economic data during this period.

3.5 Exogenous Events

The previous section identified considerable differences between the two classification techniques in terms of their interpretation of newspaper sentiment during three distinct time periods. Moreover, while newspapers are positively correlated across both textual techniques, such correlations are not very large. This phenomenon is investigated further by examining the extent to which the news sentiment indicators behave differently around historical events. A hands-off approach is considered, where structural changes in the mean of the sentiment indices are determined without any prior assumptions about the dates defining

relevant external events. For that purpose, the Bai-Perron test[27] is used, as it allows for the identification of changes in data patterns without any priors. Online Appendix B illustrates the break dates for each newspaper by applying the Bai-Perron procedure to the Naïve Bayes (Figure B.1) and Dictionary (Figure B.2) indices.

Table 9 outlines those dates identified as defining structural shifts, and shows that the procedure manages to detect a number of events that impacted on the economy over the sample period. The Bai-Perron procedure shows strong attestation of the global financial crisis by identifying the start of the global crisis meltdown (July/August 2007) in the case of 15 out of 26 economic sentiment indices (across both textual techniques). Of the newspaper sentiment indices that do not respond to the start of the financial crisis, most are Naïve Bayes indicators: such as The Daily & Sunday Mirror, Daily Record, The Sun, The Financial Times, The Sunday Telegraph and The Times. The absence of a break in the case of the last three of these newspapers is slightly surprising given their status as quality newspapers.

From observation of Table 9, it is apparent that 1998 is frequently identified by the Bai-Perron procedure; structural breaks occur for five newspapers on this year using the Dictionary method. The structural shift at this time may be due to new functions of the Monetary Policy Committee following the Bank of England Act 1998, and resulting spikes in interest rates and inflation. The year 2003, identified seven times, coincides with increased global uncertainty due to the Iraq War. Interestingly, 2012 and 2013 are associated with five and nine breaks respectively. This could be due to instability in financial markets regarding the EU Debt crisis (which is in some respects different from the global financial crisis as it involved a less concentrated timeline, where the Greek membership was considered across

[27] See Bai and Perron (2001). For each newspaper, the equation $S_t = \pi_t + \varepsilon_t$ is estimated where π_t is an intercept term. Bai-Perron sequentially tests whether $\pi_t = \bar{\pi}_r$ where $\bar{\pi}_r$ is the fixed mean for regime r.

different periods). The year 2013 also coincided with the start of discussions regarding a referendum on the UK's membership of the EU.

Table 9. Break Points Identified Using the Bai-Perron Procedure

	Naïve Bayes	Dictionary
Daily & Sunday Mail	2000 Jun, 2007 Aug, 2013 Mar	1998 May, 2007 Aug, 2012 Nov
Daily & Sunday Mirror	1999 May, 2005 Jan	2008 Mar, 2013 Feb
Daily Record	2004 Mar, 2008 May	1998 Apr, 2001 Nov, 2007 Nov
Daily Telegraph	2007 Jul, 2013 Jun	1998 May, 2007 Aug, 2012 Nov
Evening Standard	1999 Jul, 2007 Aug, 2013 Mar	2003 Aug, 2007 Aug, 2013 Apr
Financial Times	1996 Apr, 2011 Feb	2000 Aug, 2007 Aug, 2013 Mar
Guardian	1993 Dec, 2003 Aug, 2007 Aug, 2012 Dec	1994 Jan, 1998 Aug, 2003 Aug, 2007 Aug, 2013 Jan
Independent	2007 May, 2011 May	1994 Jun, 1998 Jul, 2003 Aug, 2007 Aug, 2012 Sep
Observer	1997 Mar , 2007 Aug	2007 Aug, 2013 Jan
Sun	2005 Aug	2013 Feb
Sunday Telegraph		2003 Apr, 2007 Jul, 2012 Sep
Sunday Times	2007 Apr	1994 Jan, 2003 Jul, 2007 Aug, 2011 Dec
Times	2005 Mar	2003 May, 2007 Aug

Notes: The table illustrates break dates in the sentiment time series identified by the Bai-Perron method using both Naïve Bayes and Dictionary sentiment classification techniques.

It is worth adding that the Bai-Perron procedure fails to identify the UK's EU Referendum as a break in the mean. This is as a result of a 5% sample truncation at both ends of the sample used by the procedure, and the EU Referendum period falls

within this truncated sample. Finally, it is worth noting that the highest number of events is found to occur for two quality newspapers: The Guardian and The Independent. Thus, these newspapers tend to show greater sensitivity to economic events.

3.6 Movement of Consumer and News Sentiments

Having thoroughly investigated the heterogeneity across newspapers, we extend the analysis to consumer sentiments. The next hypothesis concerns the extent to which newspaper sentiments are reflected in subsequent survey-based measures of National Economic Sentiment. To measure this effect the UK Economic Sentiment Index (ESI) is used,[28] which is a composite economic indicator incorporating five component survey-based confidence indices, and thus reflects assessments of future economic prospects from agents on both the demand and supply side of the economy. Monthly ESI values are derived using a weighted subset of 15 questions across the five surveys, with weightings assigned based on economic reasoning. Index data is collected from the European Commission website at a monthly frequency for the purposes of this study.

We estimate the following univariate regression model to test whether there is some potential for newspaper sentiment to influence subsequent movements in the survey sentiment changes. For each newspaper across textual techniques, the following univariate model is considered:

$$\Delta y_t = \alpha + \sum_{j=0}^{k} \delta_j \Delta y_{t-j} + \sum_{j=0}^{l} \beta_j s_{t-j} + \sum_{j=0}^{m} \gamma_j x_{t-j} + \varepsilon_t, \qquad (6)$$

[28] https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys/download-business-and-consumer-survey-data/time-series_en

where Δy_t is the logarithmic change of ESI, k is the number of lags for the logarithmic change of ESI, l is the number of lags for the change in the newspaper-specific sentiment index, s is the change in monthly newspaper-specific sentiment variable, x_t denotes the log-change in industrial production and price levels, ε_t represents the white noise term, and $\alpha, \delta, \gamma, \beta$ are regression coefficients.[29]

Equation (6) is an ARDL model where the logarithmic change of ESI is determined by its previous lags, lags of the changes in sentiment indices, and changes in both price levels and industrial production. To select the best model, the general to specific approach from Oxmetrics[30] is used (Krolzig and Hendry, 2001; Campos et al., 2005). A general unrestricted model (GUM) is first specified, where the maximum number of lags for *k, l* and *m* is set to 12. The model is estimated with $2^{(k+l+m+1)}$ combinations of regressors. The estimation procedure uses backward elimination along multiple paths by removing non-significant regressors each time. The removal of non-significant regressors is based on individual t-statistics, multiple hypothesis tests (parsimonious encompassing) and other diagnostic tests.[31]

After applying the reduction procedure, where multiple models are deemed to have passed the diagnostic tests, the best model is selected according to fit criterion (Bayesian Information Criterion). The statistical tests are performed at the 5% level.[32] It is noted that all the variables are stationary, which is important for inference purposes and the selection of the 'best' model. In Equation (6), the survey economic sentiments are controlled for the evolution of economic fundamentals such as business cycle or prices, by the inclusion of the log changes in industrial production and the consumer price index.

The results from estimating Equation (6) are reported in Tables D.1 (Naïve Bayes) and D.2 (Dictionary), found in Appendix D. The coefficients of the best models show strong evidence to suggest that current and previous lags of the

[29] All the model variables are used in their standardized form to facilitate interpretation of the results.
[30] An alternative package (GETS) is available in Rstudio which performs a similar function.
[31] These include tests for structural breaks, serial correlation and conditional heteroscedasticity.
[32] For a more in-depth discussion see Doornik (2009).

newspaper-based indices lead to changes in ESI. Considering Table D.1, it is found that – after controlling for changes in industrial production and inflation – a number of lags of the Naive Bayes sentiment indices influence ESI. For most of the newspapers, this effect persists for a period of four or five months. Similar results are witnessed in the case of the Dictionary technique (Table D.2). The persistence of both techniques is perhaps better illustrated in Figure 5, which plot significant coefficients reported in D.1 and D.2 and show how economic sentiment (ESI) changes respond to lagged changes in newspaper indices for both the Naïve Bayes and Dictionary techniques.

In the case of both Dictionary and Naïve Bayes techniques, the News Reaction Curves shown in Figure 5 provide evidence that the effect of news persists for a significant period. Considering quality newspapers, for both techniques the effect of news is initially strong,[33] with the greatest reaction occurring within the first three months for most newspapers before gradually decaying throughout 12 months. This effect is strongest for The Daily Telegraph and The Financial Times using the Naïve Bayes method. As for the Dictionary indices, most quality newspapers exhibit this pattern. In particular, the effect is strongest for The Evening Standard, The Sunday Telegraph, The Observer, and The Times. In all these cases, there are some lags of the change in sentiment index such that one standard deviation increase in their values, entails an increase in the standardized log change of ESI in the region of around 0.40-0.50.

[33] This trend is not observed in the case of The Observer (Naïve Bayes) , which has negative coefficients and follows a pattern inconsistent with the other quality newspapers.

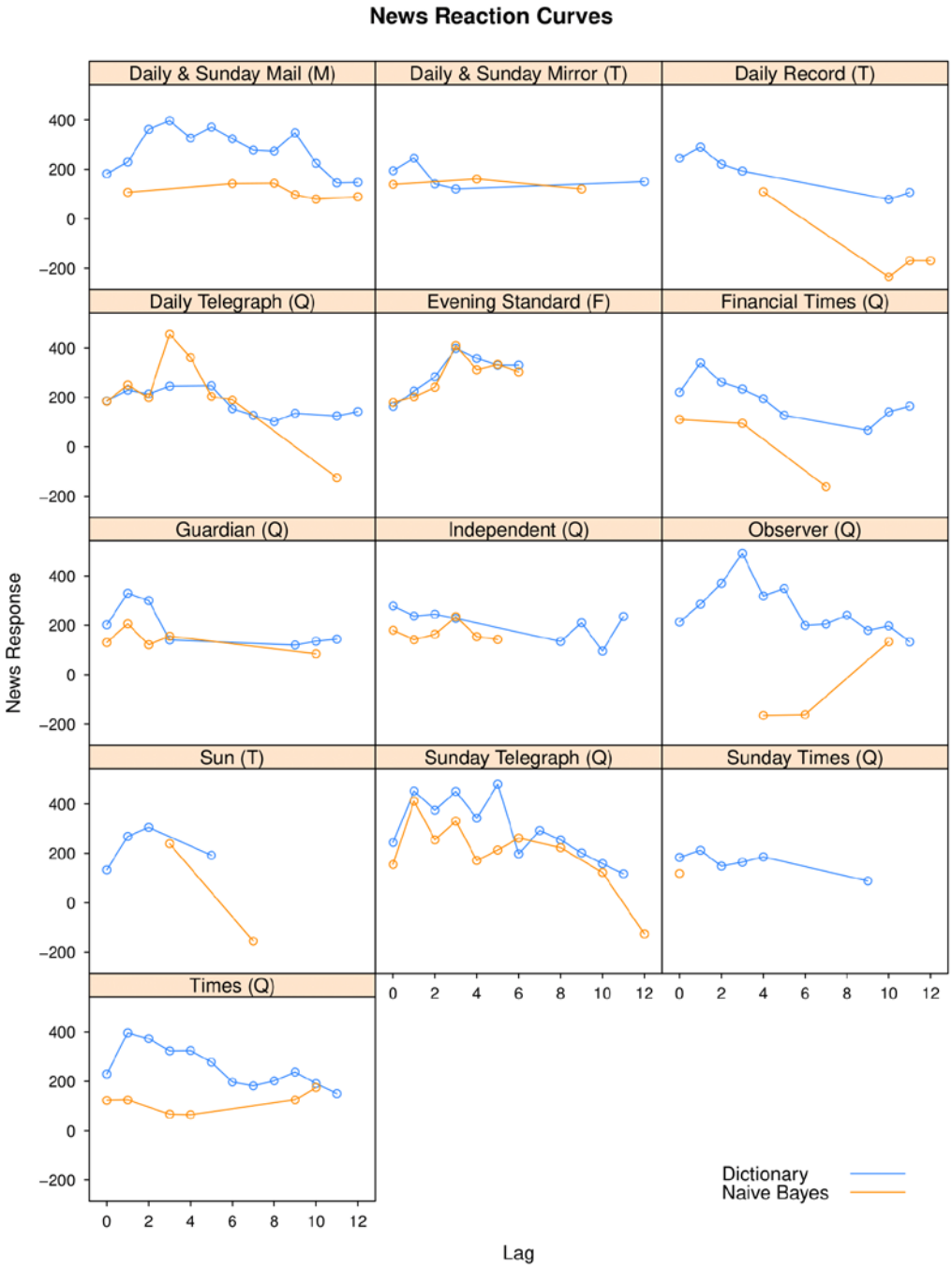

Figure 5: News reaction lags for quality and tabloid UK newspapers using the Naïve Bayes and Dictionary techniques. The figure shows the value of statistically-significant standardized parameters (multiplied by 1,000) for each lag of change of a news-based sentiment index, based on the optimal model for Equation (6). (Q) denotes that the newspaper is considered to be a quality (broadsheet) publication format, as indicated in Section 2.1, while (T) indicates a tabloid newspaper format, (M) represents a middle-market publication and (F) indicates a free newspaper.

For both Dictionary and Naïve Bayes techniques, the News Reaction Curves provide evidence that the effect of news persists for a significant period. In the case of quality newspapers, the effect of news is initially strong using both techniques,[34] with the greatest reaction occurring within the first three months for most newspapers before gradually decaying throughout 12 months. This effect is particularly strong for The Daily Telegraph and The Financial Times using the Naïve Bayes method, while for Dictionary indices it is most quality newspapers that exhibit this pattern. Tabloid newspapers such as The Sun, The Daily Record and The Daily & Sunday Mirror don't tend to fare well, and their impact is relatively low in comparison to broadsheet publications. The Dictionary indices show muted increases of between 0.10 and 0.25, and these changes tend to remain only up to a maximum of three months. For Naïve Bayes indices the impact is less obvious, possibly suggesting a longer persistence of seven to nine months, but with negative coefficients of between -0.10 and -0.20 at higher lags in some cases. The poor effect of tabloids is rather expected in this context given that the coverage of economic news is limited in these newspapers, which implies a weaker effect on overall economic sentiments.

3.7 Out of Sample Forecast Accuracy

This section compares the news-based sentiment indicators in terms of forecast performance. We construct and evaluate point forecasts for the changes in economic sentiment indices using four unique models which incorporate the different newspaper indicators. Considering that the economic sentiment indicators already have predictive power, we only examine the forecasting power of the news sentiment index towards forecasting the economic sentiment indicators instead of macroeconomic indicators.

Forecasting performance is intended to illustrate that the news-based sentiment indicators can be used for forecasting purposes. The models considered are simple; taking the form of an autoregressive model with the lagged sentiment index as a regressor and VAR(p) models:

$$\Delta y_t = \alpha + \sum_{j=0}^{k} \delta_j \Delta y_{t-j} + \beta_1 s_{t-1} + \varepsilon_t, \quad (7)$$

$$Z_t = B_0 + B_1 Z_{t-1} + \cdots + B_p Z_{t-p} + u_t, \quad (8)$$

where Z_t in Equation (8) is a vector containing changes in the economic sentiment index, news sentiment index, industrial production and price levels. This is compared to an autoregressive process as a benchmark:

$$\Delta y_t = \alpha + \sum_{j=0}^{k} \delta_j \Delta y_{t-j} + \varepsilon_t. \quad (9)$$

To test for predictive power, we estimate the models and forecast changes in the Economic Sentiment Index for an out-of-sample period of *h =1-12* months. The forecasts are constructed using a rolling window scheme, where parameters are estimated using the latest 60 observations, and then used to forecast ahead. Forecasts are produced for the period spanning January 2005 until December 2016. For each horizon, the average root mean squared error is computed. With respect to the number of lags in the model, Table 10 reports the results for *k=1, 3* and *6*.[35]

Conditional on econometric specification, newspapers and textual method, a model is considered successful if the root mean squared error is lower than that of the benchmark. For instance, in Panel A, the figures show the percentage number of observations for which all the models estimated outperform the benchmark for the specific horizon (*h*). *k* refers to the number of lags included as autoregressive terms in both Equations (7) and (8).

[35] Lags of nine and twelve months were found to be rather weak in the forecasting process.

Table 10. Forecasting Power of News-Based Sentiment Indicators

Panel A: Performance across time horizons (January 2005-December 2016)

	***h*=**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
k =1		26.9	71.2	69.2	61.5	63.5	63.5	57.7	57.7	67.3	55.8	17.3	9.6
k =3		9.6	69.2	69.2	61.5	57.7	61.5	63.5	57.7	63.5	57.7	73.1	57.7
k =6		5.8	51.9	65.4	44.2	59.6	61.5	57.7	53.8	61.5	51.9	73.1	59.6

Panel B: Performance across Textual method and Specification

			ARDL		**VAR**	
	ARDL	**VAR**	**DIC**	**NB**	**DIC**	**NB**
k =1	26.3	77.2	23.7	28.8	76.9	77.6
k =3	24.7	92.3	19.9	29.5	92.9	91.7
k =6	20.8	86.9	17.9	23.7	86.5	77.6

Panel C: Performance across Newspapers

	k =1				**k =3**				**k =6**			
	$ARDL_{NB}$	$ARDL_{DIC}$	VAR_{NB}	VAR_{DIC}	$ARDL_{NB}$	$ARDL_{DIC}$	VAR_{NB}	VAR_{DIC}	$ARDL_{NB}$	$ARDL_{DIC}$	VAR_{NB}	VAR_{DIC}
Daily & Sunday Mail	8	8	75	75	25	8	92	92	25	17	83	92
Daily & Sunday Mirror	33	0	75	75	0	17	92	92	17	8	92	83
Daily Record	25	8	75	75	8	8	92	92	0	0	92	92
Daily Telegraph	33	33	75	83	50	25	92	92	50	33	92	75
Evening Standard	33	50	75	75	42	8	92	100	33	17	92	100
Financial Times	8	25	75	83	33	25	92	92	17	25	92	75
Guardian	33	25	83	75	58	25	92	92	42	17	83	92
Independent	58	33	75	75	58	25	92	92	33	17	92	92
Observer	25	0	83	75	8	8	92	92	8	25	92	83
Sun	42	17	75	75	42	25	92	100	33	17	92	92
Sunday Telegraph	17	25	83	75	25	42	92	92	17	25	92	92
Sunday Times	33	42	75	83	8	25	92	92	0	25	58	83
Times	25	42	75	83	25	17	92	92	33	8	83	75

Notes: Panel A of the figures shows the percentage of observations for which all models outperform the benchmark. Panel B shows the percentage of observations that the textual specific models and econometric specifications outperform the benchmark. Panel C illustrates the performance of each newspaper across textual method and econometric specification. The sub-headings refer to equations 7 and 8, and NB and DIC refer to the Naïve Bayes and Dictionary method.

In a nutshell, much of the performance depends on econometric specification used. The VAR model is comparatively more successful than the autoregressive model. From Panel A of Table 10, forecasting performance appears to be strong in the medium term (between three and nine months ahead). Forecasts for 11 months ahead seem to be particularly accurate for *k =3* and *6*, while performance is found to be poor for *k=1*. Forecasting for three months ahead shows few differences across different lag periods. Panel B shows that the VAR model outperforms the univariate specification considerably: the percentage success rate for VAR is above 75% while for the univariate model it is between 20-30%. There are no major differences between the Naïve Bayes and Dictionary models when used in conjunction with the VAR specification whereas the Naïve Bayes technique slightly outperforms the Dictionary in the univariate model.

Panel C shows individual newspaper performance across the different techniques. Although no evidence of major differences in performance is identified for the VAR models, considerably more disparity is observed for the univariate models. In this category, models with lags k=3 and k=6 appear to be more successful in the forecasting performance. Good forecasters that are consistent across lag periods, textual techniques, and model specification include The Daily Telegraph, The Evening Standard and The Guardian. Though not so consistent throughout, The Times and The Sunday Times are also found to perform particularly well.

4. Conclusion

This study establishes new measures of economic sentiment by applying two distinct textual analysis techniques to a large dataset of economic policy-oriented news articles. Specifically, articles published within thirteen nationally-distributed newspapers in the UK are used to construct time series of newspaper sentiment over

time. In doing so, we provide novel evidence concerning the coverage of economic news in printed media. The main body of the analysis focus on the degree of heterogeneity in newspapers' portrayal of economic news: in terms of newspaper-specific tonality (sentiment), the dispersion of tonality within each newspaper (as a measure of consensus), and the intensity (frequency) of relevant articles published. The usefulness of newspaper sentiment indices for forecasting future changes in the European Sentiment Index (ESI) is also assessed.

We find that newspapers differ to various degrees in the way that economic policy news is reported. In particular, a clear divide is identified between two distinct newspaper formats: quality (broadsheet) and tabloid publications. Demographic profiles of quality and tabloid newspaper readership suggests that both publication formats tend to appeal to different socio-economic and demographic categories. As such, our findings support the suggestion that publications may portray news in context which appeals to a certain readership, and thus maximises their own profits (Gentzkow and Shapiro, 2010; Mullainathan and Shleifer, 2005).

Previous research has found that newspaper coverage of environmental issues differs based on the format of the publication (Boykoff, 2008), and we find this trend to extend to economic reporting also: newspaper publication format shapes the way in which policy-related economic news is communicated. Quality newspapers are found to report much frequently on UK economic policy-related news, and economic sentiment portrayed in such publications tends be objective (unbiased) in nature. On the other hand, tabloids are found to publish much less frequently on topics relating to economic policy, and coverage in these publications tends to portray more negative economic sentiments.

The nature of newspaper coverage has been shown to change around times of increased economic uncertainty in the United States (Baker et al., 2016). We therefore focus on two specific exogenous events – the global financial crisis and the period following the UK's EU Referendum – to provide additional insights

regarding the portrayal of economic news in UK newspapers during times of increased uncertainty. Notable decreases in economic sentiment are detected within most UK newspapers during both periods. We find that quality newspapers tend to publish more than their tabloid counterparts during times of extreme uncertainty, although tabloids do witness substantial increases in frequency compared to the baseline during both events.

Sentiment precision (or consensus) across news articles add further insight to the way information is communicated. Whereas a reduction of consensus between newspapers is identified during the global financial crisis, the period following the EU Referendum exhibits stronger consensus, albeit with a greater variance across newspapers. The evidence points to the possibility of some newspapers taking an entrenched position around the implications of the EU Referendum vote while others may be communicating a more balanced range of viewpoints on the topic.

Evidence of a news reaction curve is found, where our results point towards newspapers having lagged effects on consumer sentiments. Thus, our findings add further credence to the observation that news intensity of economic factors tends to be self-fulfilling (Lamla and Lein, 2014), as is the intensity with which economic-policy related news is reported. The two textual techniques (Naïve Bayes and Dictionary) display different memory patterns: Dictionary indices appear to exhibit higher lags on consumer sentiment over time, whereas for Naïve Bayes indices the number of lags is smaller, and the number of newspapers fewer.

Finally, the application of the derived newspaper sentiment time series in forecasting future consumer sentiment is considered. It is found that newspapers combined with a Vector Autoregression outperform the benchmark of a simple autoregressive model. A simple autoregression with the sentiment variable also outperforms the benchmark, although the improvement in this case is modest. Thus, our constructed economic indicators appear to have some use for effective forecasting of consumer sentiment survey outcomes.

This paper does not directly address the causes of observed differences between newspapers in the way in which economic news is reported, but some observations can be made to shed light on this issue. Differences are likely to occur given that newspapers have different target audiences (Gentzkow and Shapiro, 2010). Thus, newspapers may report economic news in a context that satisfies the expectations of their consumers. For example, quality newspapers are more likely to command a greater readership within higher socioeconomic categories, and to meet the expectations of this readership they are more likely to (i) publish more news content relating to economic policy and (ii) be more objective in their coverage.

References

Baker, Scott. R., Bloom Nicholas, and Davis, Steven J., 2016. Measuring economic policy uncertainty. The Quarterly Journal of Economics 131(4), 1593-1636.

Blinder, Alan S., Krueger, Alan B., 2004. What does the public know about economic policy, and how does it know it? Brooking Papers on Economic Activity 1, 327-387.

Boykoff, Maxwell T., 2008. The cultural politics of climate change discourse in UK tabloids. Political Geography 27(5), 549-569.

Brown, Alasdair, Rambaccussing, Dooruj, Reade, James J., Rossi, Giambattista, 2018. Forecasting with social media: Evidence from tweets on soccer matches. Economic Inquiry 56(3), 1748-1763.

Campos, Julia, Ericsson, Neil R., Hendry, David F., 2005. General-to-Specific Modelling: an Overview and Selected Bibliography. Board of Governors of the Financial Reserve System, International Finance Discussion Papers 838, 1-92.

DellaVigna, Stefano, Kaplan, Ethan, 2007. The Fox News effect: Media bias and voting. The Quarterly Journal of Economics 122(3), 1187-1234.

Doms, Mark E., Morin, Norman J., 2004. Consumer Sentiment, the Economy, and the News Media. Working Paper. Federal Reserve Board of Governors.

Gentzkow, Matthew, Shapiro, Jesse M., 2010. What drives media slant? Evidence from US daily newspapers. Econometrica 78(1), 35-71.

Gentzkow, Matthew, Kelly, Byan T., Taddy, Matt, 2017. Text as Data. Working Paper. National Bureau of Economic Research.

Hamilton, James, 2004. All the news that's fit to sell: How the market transforms information into news. Princeton University Press, Woodstock.

Hetherington, Marc J., 1996. The media's role in forming voters' national economic evaluations in 1992. American Journal of Political Science 40(2), 372-395.

Kohavi, Ron, Provost, Foster, 1998. Guest editors' introduction: On applied research in machine learning. Machine Learning, 30(2-3), 271-274.

Koppel, Moshe, Shtrimberg, Itai, 2006. Good news or bad news? Let the market decide. In: Shanahan, J.G., Qu, Y., Wiebe, J. (Eds.), Computing attitude and affect in text: Theory and applications. Springer, Dordrecht, pp. 297-301.

Krolzig, Hans-Martin, Hendry, David F., 2001. Computer automation of general-to-specific model selection procedures. Journal of Economic Dynamics and Control 25(6-7), 831-866.

Lamla, Michael J., Lein, Sarah M., 2014. The role of media for consumers' inflation expectation formation. Journal of Economic Behavior & Organization 106, 62-77.

Lamla, Michael J., Maag, Thomas, 2012. The role of media for inflation forecast disagreement of households and professional forecasters. Journal of Money, Credit and Banking 44(7), 1325-1350.

Lamla, Michael J., Lein, Sarah M., 2015. Information rigidities, inflation perceptions, and the media: Lessons from the Euro cash changeover. Economic Inquiry 53, 9–22.

Ludvigson, Sydney C., 2004. Consumer confidence and consumer spending. Journal of Economic Perspectives 18 (2), 29-50.

Mullainathan, Sendhil, Shleifer, Andrei, 2005. The market for news. The American Economic Review 95(4), 1031-1053.

Starr, Martha A., 2012. Consumption, sentiment, and economic news. Economic Inquiry 50(4), 1097-1111.

Tobback, Ellen, Naudts, Hans, Daelemans, Walter, de Fortuny, Enric J., Martens, David, 2016. Belgian economic policy uncertainty index: improvement through text mining. International Journal of Forecasting 34(2), 355-365.

Tetlock, Paul C., 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance 62(3), 1139-1168.

The Economist, 2011. The R-word index: Gauging the gloom. The Economist. Available at https://www.economist.com/graphic-detail/2011/09/16/gauging-the-gloom.

Wang, Xiao-Wei, Nie, Dan, Lu, Bao-Liang, 2014. Emotional state classification from EEG data using machine learning approach. Neurocomputing 129, 94-106.

Appendix A. UK Newspaper Readership: Descriptive Statistics

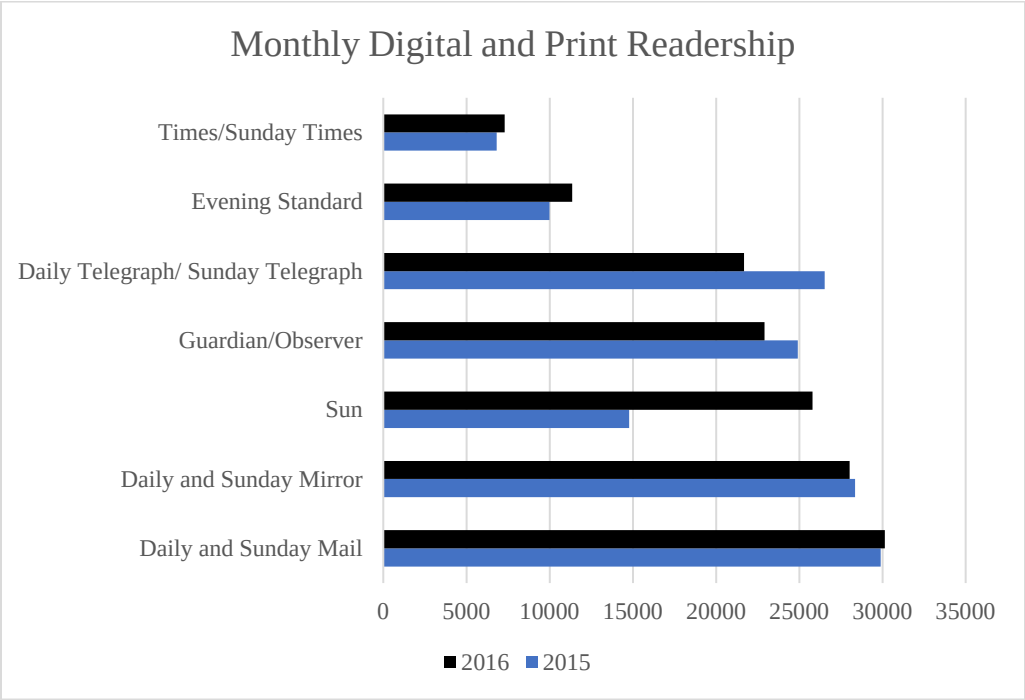

Figure A.1. Monthly Digital and Print Readership. The bar chart shows the number of readers (in thousands) for each newspaper in our sample, based on Ofcom data. The readership is presented for 2015 and 2016. Newspapers are shown on the vertical axis and the number of readers are shown on the horizontal axis.

Table A.1. Demographic Properties

	All	Male	Female	16-34	65+	ABC1	C2D2
Daily & Sunday Mail	41	41	40	33	43	41	39
Sun	32	38	26	44	27	21	45
Daily & Sunday Mirror	17	19	14	21	20	12	23
Guardian	9	8	9	18	3	12	5
Times	9	9	9	10	8	15	2
Sunday Times	9	10	9	11	6	15	3
Daily Telegraph	7	7	8	12	10	10	4
Evening Standard	5	5	6	1	6	6	2
Independent	5	3	6	1	6	6	2
Sunday Telegraph	5	5	4	4	6	8	1
Observer	4	3	5	3	3	6	2
Financial Times	2	2	2	8	1	4	0
Daily Record	2	2	2	2	2	1	2

Notes: The table shows the percentage of respondents (from an Ofcom survey) who use the newspaper listed in the first column for news consumption. The respondents are categorised based on age group (above 16, 16-34 and 65+), gender and socio-economic status, where ABC1 represents the highest socioeconomic class, and C2D2 represents the lowest. The original Ofcom data includes other newspapers not examined in this study. The table should be read as follows: For the Daily & Sunday Mail, 41% of all respondents use this publication as a source of news. Similarly, 41% of male respondents and 40% of female respondents read The Daily & Sunday Mail

Appendix B. Correlations across News-Based Indices

Table B.1. Correlations across News-Based Indices (Naive Bayes Technique: November 2000 - December 2016)

	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
Daily & Sunday Mail	1.000												
Daily & Sunday Mirror	0.409	1.000											
Daily Record	0.238	0.285	1.000										
Daily Telegraph	0.614	0.396	0.327	1.000									
Evening Standard	0.634	0.491	0.378	0.607	1.000								
Financial Times	0.390	0.367	0.250	0.511	0.460	1.000							
Guardian	0.605	0.361	0.323	0.636	0.615	0.476	1.000						
Independent	0.491	0.470	0.414	0.538	0.630	0.605	0.514	1.000					
Observer	0.364	0.382	0.227	0.350	0.383	0.242	0.361	0.395	1.000				
Sun	0.422	0.394	0.266	0.425	0.440	0.440	0.418	0.473	0.302	1.000			
Sunday Telegraph	0.325	0.104	0.108	0.331	0.290	0.275	0.300	0.275	0.162	0.251	1.000		
Sunday Times	0.422	0.293	0.188	0.426	0.464	0.357	0.407	0.416	0.389	0.240	0.169	1.000	
Times	0.538	0.463	0.250	0.596	0.602	0.499	0.430	0.611	0.442	0.381	0.299	0.421	1.000

Notes: Table of correlation across the news-based index using the Naïve Bayes method for the period November 2000 until December 2016.

Table B.2. Correlations across News-Based Indices (Dictionary Technique: November 2000 - December 2016)

	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
Daily & Sunday Mail	1.000												
Daily & Sunday Mirror	0.503	1.000											
Daily Record	0.489	0.518	1.000										
Daily Telegraph	0.757	0.497	0.480	1.000									
Evening Standard	0.722	0.449	0.397	0.758	1.000								
Financial Times	0.701	0.486	0.495	0.784	0.697	1.000							
Guardian	0.732	0.525	0.493	0.776	0.716	0.725	1.000						
Independent	0.735	0.517	0.493	0.693	0.691	0.709	0.678	1.000					
Observer	0.550	0.452	0.393	0.562	0.549	0.627	0.529	0.533	1.000				
Sun	0.457	0.292	0.341	0.489	0.483	0.359	0.529	0.403	0.277	1.000			
Sunday Telegraph	0.508	0.364	0.320	0.521	0.545	0.498	0.528	0.542	0.333	0.365	1.000		
Sunday Times	0.594	0.462	0.425	0.583	0.541	0.618	0.588	0.561	0.492	0.292	0.428	1.000	
Times	0.700	0.485	0.477	0.793	0.708	0.806	0.734	0.684	0.529	0.444	0.556	0.610	1.000

Notes: Table of correlation across the news-based index using the Dictionary method for the period November 2000 until December 2016.

Appendix C. Newspaper Fixed Effects

Table C.1. Newspaper Fixed Effects: Test of Coefficient Differences (Naïve Bayes)

	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
Daily & Sunday Mail	0.000												
Daily & Sunday Mirror	0.000												
Daily Record	0.000	0.000											
Daily Telegraph	0.000	0.000	0.000										
Evening Standard	0.000	0.000	0.000	0.000									
Financial Times	0.000	0.000	0.000	0.000	0.000								
Guardian	0.000	0.000	0.000	0.000	0.000	0.000							
Independent	0.000	0.000	0.000	**0.117**	0.000	0.000	0.000						
Observer	**0.511**	0.000	0.000	0.000	0.000	0.000	0.000	0.000					
Sun	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000				
Sunday Telegraph	0.000	0.000	0.000	0.046	0.197	0.000	0.000	**0.214**	0.000	0.000			
Sunday Times	0.000	0.000	0.000	0.010	0.000	0.000	0.017	0.000	0.000	0.000	0.001		
Times	0.000	0.000	0.000	0.000	0.096	0.000	0.000	0.000	0.000	0.000	**0.142**	0.000	

Notes: The table reports the p-value from testing the newspapers in the top row against the newspapers in the first column, testing for different fixed effects across newspapers from Equation (2). Figures in bold are cases when the null hypothesis is not rejected at the 10% significance level.

Table C.2. Newspaper Fixed Effects: Test of Coefficient Differences (Dictionary)

	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
Daily & Sunday Mail													
Daily & Sunday Mirror	**0.133**												
Daily Record	0.000	0.000											
Daily Telegraph	0.000	0.000	0.000										
Evening Standard	**0.351**	0.038	0.000	0.000									
Financial Times	0.030	0.004	0.000	0.000	**0.477**								
Guardian	0.124	**0.511**	0.000	0.000	0.018	0.000							
Independent	0.000	0.000	0.000	0.000	0.001	0.000	0.000						
Observer	0.000	0.000	0.000	**0.205**	0.000	0.001	0.000	**0.318**					
Sun	0.000	0.000	0.000	0.480	0.000	0.000	0.000	0.002	**0.123**				
Sunday Telegraph	0.000	0.000	0.028	0.000	0.000	0.000	0.000	0.000	0.000	0.000			
Sunday Times	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.369**		
Times	0.000	0.000	0.000	0.021	0.000	0.000	0.000	0.000	0.008	**0.526**	0.000	0.000	

Notes: The table reports the p-value from testing the newspapers in the top row against the newspapers in the first column, testing for different fixed effects across newspapers from Equation (2). Figures in bold are cases when the null hypothesis is not rejected at the 10% significance level.

Appendix D. ARDL Model Estimated Parameters

Table D.1. Estimated Parameters for Equation (6) with Naïve Bayes Method

Lag of x	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
0		139.21*		184.27*	178.95**	110.58.	130.66*	180.26**			154.84*	118.25*	122.64*
1	107.05*			249.63.	201.65**		206.11***	141.94*			411.68***		124.32*
2				198.66	241.02**		121.73.	163.76.			254.67**		
3				456.21***	409.27***	94.98.	155.72*	235.01*		238.99***	330.59***		66.65
4		161.05**	109.40*	361.58**	311.60***			154.05.	-164.48***		171.16*		64.52
5				203.45	333.02***			142.99*			213.33***		
6	142.33**			189.66*	303.18***				-161.32**		261.77***		
7						-161.55*				-155.75*			
8	143.77*										223.00***		
9	97.38.	120.74*											124.76*
10	79.83		-235.25***				85.38.		133.68**		122.43*		175.41***
11			-169.52*	-125.56*									
12	89.25		-169.68*								-125.97*		
R-squared	0.237	0.246	0.237	0.390	0.255	0.207	0.203	0.211	0.257	0.358	0.412	0.160	0.226
Log-lik.	-368.12	-316.56	-336.90	-214.71	-359.73	-374.66	-404.01	-406.17	-342.95	-224.58	-215.97	-418.75	-405.77
BIC	855.38	699.18	751.82	549.11	815.66	828.75	899.86	927.33	764.33	543.22	541.68	895.06	926.65
AIC	778.24	657.12	701.81	475.42	753.45	777.33	840.03	852.34	713.90	485.16	473.93	857.50	851.54

Notes: The table illustrates the results from applying the General to Specific Modelling approach using the Naive Bayes technique (standardized coefficients). The regression coefficients are multiplied by 1000. *, ** and *** indicates statistical significance at the 10%, 5% and 1%.

Table D.2. Estimated Parameters for Equation (6) with Dictionary Method

Lag of x	Daily & Sunday Mail	Daily & Sunday Mirror	Daily Record	Daily Telegraph	Evening Standard	Financial Times	Guardian	Independent	Observer	Sun	Sunday Telegraph	Sunday Times	Times
0	182.44*	194.41**	245.49***	185.70**	162.99**	220.96***	202.08**	278.45***	213.23**	133.22	243.84**	183.44**	229.31***
1	229.57**	245.72**	290.29***	229.65**	225.46***	339.09***	329.81***	237.34***	285.77**	268.21**	452.06***	212.08*	396.40***
2	362.42***	141.38	220.83**	213.81***	283.33***	261.61**	300.37***	245.01***	369.30***	304.98***	374.83***	148.36*	372.86***
3	397.24***	120.86.	193.31**	245.05***	398.44***	233.74**	140.83*	227.84***	491.69***		450.26***	164.17*	322.88***
4	326.94***				357.63***	194.35**			318.78***		342.40***	185.56**	324.37***
5	371.59***			246.98***	330.27***	127.37.			348.89***	192.01**	479.46***		277.25***
6	324.14***			153.79.	251.40***				199.54*		197.78*		197.19**
7	278.60***			126.87.					204.95*		292.10***		181.69*
8	274.66**			101.86				134.75*	241.14**		254.48**		201.95**
9	348.58***			135.36*		66.86	120.48*	210.20***	179.28*		201.59**	88.02.	237.39***
10	224.47**		78.88			141.09.	135.81*	96.48.	197.75*		158.22.		190.77**
11	145.43*		106.82	124.69*		164.26**	144.13**	235.70***	132.97*		117.17*		150.44*
12	147.55**	150.32**		142.07*									
R-squared	0.248	0.272	0.273	0.421	0.269	0.262	0.280	0.301	0.303	0.390	0.438	0.193	0.258
Log-lik.	-365.96	-312.14	-330.54	-209.96	-356.96	-364.29	-388.14	-387.14	-334.40	-219.87	-211.72	-412.47	-399.16
BIC	873.76	717.87	761.38	550.03	815.80	836.37	914.03	895.02	803.25	539.04	569.76	911.27	919.18
AIC	781.93	658.28	697.09	469.93	749.93	766.58	824.28	816.28	716.80	477.75	479.44	854.94	840.31

Notes: The table illustrates the results from applying the General to Specific Modelling approach using the Dictionary technique (standardized coefficients). The regression coefficients are multiplied by 1000. *, ** and *** indicates statistical significance at the 10%, 5% and 1%.